AMENDMENTS TO THE CLAIMS

Please enter the following amendments without prejudice or disclaimer.

Please cancel claims 9-11, 13, and 18-20 without prejudice or disclaimer.

This listing of claims will replace all prior versions, and listings, of claims in the application:

In the claims:

Claim 1 (previously presented): The method of claim 34, wherein the property of interest is a target for a drug.

Claim 2 (previously presented): The method of claim 34, wherein the property of interest is that of being essential for the growth or viability of an organism.

Claim 3 (previously presented): The method of claim 1, wherein the drug is an anti-microbial drug.

Claim 4 (previously presented): The method of claim 1 or claim 2, wherein the first nucleic acid sequence or polypeptide sequence is derived from a pathogen.

Claim 5 (original): The method of claim 4, wherein the pathogen is a microorganism.

Claim 6 (previously presented): The method of claim 5, wherein the microorganism is Mycobacterium tuberculosis (MTB).

Claim 7 (original): The method of claim 1 or claim 2, wherein the plurality of sequences used to identify a second sequence comprises a database of the gene sequences of an entire genome of an organism.

Claim 8 (original): The method of claim 1 or claim 2, wherein the plurality of sequences used to identify a second sequence comprises a database of the gene sequences derived from a pathogen.

Claim 9-11 (canceled)

Claim 12 (currently amended): A method for identifying a second nucleic acid sequence or second polypeptide sequence of a second protein, wherein the second protein has a biological or chemical property of interest, comprising:

(a) providing a first nucleic acid sequence that encodes a first protein, or a first polypeptide sequence of the first protein, wherein the first protein has a biological or chemical property of interest;

(b) providing an algorithm capable of analyzing a functional relationship between the first protein and second protein, wherein the algorithm is a "phylogenetic profile" method, wherein the "phylogenetic profile" method algorithm comprises

(i) obtaining data comprising a plurality of sequences, wherein the plurality of sequences comprises a list of polypeptide sequences of proteins from at least two genomes or a list of nucleic acid sequences that encode proteins from at least two genomes;

(ii) determining a protein phylogenetic profile for the first protein and for each protein of the plurality of sequences, wherein the protein phylogenetic profile indicates the presence or absence of a protein belonging to a particular protein family in each of the at least two genomes wherein the presence or absence of a protein in a particular protein family is determined by homology,

wherein the homology between proteins is considered significant if a probability (p) of obtaining a higher homology score when the sequences are shuffled is below a probability (p) value threshold and ~~The method of claim 11,~~ wherein the probability (p) value threshold is set with respect to the value 1/NM, based on the total number of sequence comparisons that are to be performed, wherein N is the number of proteins in the first organism's genome and M is the number of proteins in all other genomes;

(iii)grouping the proteins of the plurality of sequences based on similar profiles, wherein proteins with similar profiles are indicated to have a functional relationship; and

(iv)comparing the first nucleic acid sequence or the first polypeptide sequence to the plurality of sequences by comparing the protein phylogenetic profile for the first protein to the protein phylogenetic profiles of the plurality of sequences to identify the second protein, whereby the second protein is selected from the members of the group with similar profiles as the first protein; and

(c) comparing the first nucleic acid sequence or the first polypeptide sequence to a plurality of sequences using the algorithm as set forth in step (b) to identify the second nucleic acid sequence or second polypeptide sequence of the second protein which has a functional relationship to the first protein; thereby identifying a second nucleic acid sequence or a second polypeptide sequence of a second protein that possesses the property of interest.

Claim 13 (canceled)

Claim 14 (currently amended): A method for identifying a second nucleic acid sequence or second polypeptide sequence of a second protein, wherein the second protein has a biological or chemical property of interest, comprising:

(a) providing a first nucleic acid sequence that encodes a first protein, or a first polypeptide sequence of the first protein, wherein the first protein has a biological or chemical property of interest;

(b) providing an algorithm capable of analyzing a functional relationship between the first protein and second protein, wherein the algorithm is a "phylogenetic profile" method, wherein the "phylogenetic profile" method algorithm comprises

(i) obtaining data comprising a plurality of sequences, wherein the plurality of sequences comprises a list of polypeptide sequences of proteins from at least two genomes or a list of nucleic acid sequences that encode proteins from at least two genomes;

(ii) determining a protein phylogenetic profile for the first protein and for each protein of the plurality of sequences, wherein the protein phylogenetic profile indicates the

presence or absence of a protein belonging to a particular protein family in each of the at least two genomes wherein the presence or absence of a protein in a particular protein family is determined by calculating an evolutionary distance ~~The method of claim 13, wherein the evolutionary distance is calculated~~ by:

(A) aligning two sequences from the list of proteins;

(B) determining an evolution probability process by constructing a conditional probability matrix: p(aa→aa'), where aa and aa' are any amino acids, said conditional probability matrix being constructed by converting an amino acid substitution matrix from a log odds matrix to said conditional probability matrix;

(C) accounting for an observed alignment of the constructed conditional probability matrix by taking the product of the conditional probabilities for each aligned pair during the alignment of the two sequences, represented by $P(p)=\prod_n p(aa_n \rightarrow aa'_n)$; and

(D) determining an evolutionary distance α from powers equation $p'=p^{\alpha}(aa\rightarrow aa')$, maximizing for P;

(iii) grouping the proteins of the plurality of sequences based on similar profiles, wherein proteins with similar profiles are indicated to have a functional relationship; and

(iv)comparing the first nucleic acid sequence or the first polypeptide sequence to the plurality of sequences by comparing the protein phylogenetic profile for the first protein to the protein phylogenetic profiles of the plurality of sequences to identify the second protein, whereby the second protein is selected from the members of the group with similar profiles as the first protein; and

(c) comparing the first nucleic acid sequence or the first polypeptide sequence to a plurality of sequences using at least one of the algorithms as set forth in step (b) to identify the second nucleic acid sequence or second polypeptide sequence of the second protein which has a functional relationship to the first protein, thereby identifying a second nucleic acid sequence or a second polypeptide sequence of a second protein that possesses the property of interest.

Claim 15 (original): The method of claim 14, wherein the conditional probability matrix is defined by a Markov process with substitution rates, over a fixed time interval.

Claim 16 (original): The method of claim 14, where the conversion from an amino acid substitution matrix to a conditional probability matrix is represented by:

$$P_B(i \rightarrow j) = p(j)2^{\wedge}\frac{\text{BLOSUM62}_{ij}}{2},$$

where BLOSUM62 is an amino acid substitution matrix, and P(i→j) is the probability that amino acid *i* is replaced by amino acid *j* through point mutations according to BLOSUM62 scores.

Claim 17 (original): The method of claim 16, where *Pj*'s are the abundances of amino acid *j* and are computed by solving a plurality of linear equations given by the normalization condition that:

$$\sum_i P_B(i \rightarrow j) = 1.$$

Claim 18-21 (canceled)

Claim 22 (previously presented): The method of claim 36, wherein the aligning is performed by an algorithm selected from the group consisting of a Smith-Waterman algorithm, Needleman-Wunsch algorithm, a BLAST algorithm, a FASTA algorithm, and a PSI-BLAST algorithm.

Claim 23 (previously presented): The method of claim 36, wherein at least one polypeptide sequence is obtained by translating a nucleic acid sequence from a genome database.

Claim 24 (previously presented): The method of claim 36, wherein the polypeptide or nucleic acid sequences of at least the first, second or third protein are from a database.

Claim 25 (previously presented): The method of claim 36, wherein at least the first protein has a known function.

Claim 26 (previously presented): The method of claim 36, wherein at least one of the proteins has an unknown function.

Claim 27 (previously presented) The method of claim 36, wherein the alignment is based on the degree of homology of the nucleic acid or polypeptide sequences of the first and second proteins to a segment of the nucleic acid or polypeptide sequence of the third protein.

Claim 28 (previously presented) The method of claim 36, wherein the homology between the sequences of the first and third protein and the second and third protein is considered significant if the probability (p) of obtaining a higher homology score when the sequences are shuffled is below a probability (p) value threshold.

Claim 29 (previously presented) The method of claim 28, wherein the probability (p) value threshold is set with respect to the value 1/NM, based on the total number of sequence comparisons that are to be performed, wherein N is the number of proteins in a first organism's genome and M is the number of proteins in all other genomes.

Claim 30 (previously presented): The method of claim 36, further comprising filtering excessive functional links between the first protein and any second protein.

Claims 31 to 33 (canceled)

Claim 34 (currently amended): A method for identifying a second nucleic acid sequence or second polypeptide sequence of a second protein, wherein the second protein has a biological or chemical property of interest, comprising:

(a) providing a first nucleic acid sequence that encodes a first protein, or a first polypeptide sequence of the first protein, wherein the first protein has a biological or chemical property of interest;

(b) providing ~~at least one~~ an algorithm capable of analyzing a functional relationship between the first protein and second protein, wherein the algorithm is ~~selected from the group~~

~~consisting of~~ a "domain fusion" ~~method, a "phylogenetic profile" method, and a "physiologic linkage"~~ method; and

(c) comparing the first nucleic acid sequence or the first polypeptide sequence to a plurality of sequences using ~~at least one of~~ the algorithm ~~algorithms~~ as set forth in step (b) to identify the second nucleic acid sequence or second polypeptide sequence of the second protein which has a functional relationship to the first protein, thereby identifying a second nucleic acid sequence or a second polypeptide sequence of a second protein that possesses the property of interest.

Claim 35 (previously presented): The method of claim 34, wherein the property of interest is a binding or catalytic site or cellular localization.

Claim 36 (previously presented) The method of claim 1 or 2, wherein the "domain fusion" method comprises:

(a) providing a pair of non-homologous nucleic acid or polypeptide sequences of the first and second proteins, respectively;

(b) providing a third nucleic acid or polypeptide sequence of a third protein;

(c) aligning the sequences of the first and second proteins in step (a) to a segment of the sequence in step (b); and

(d) establishing whether the first and second proteins in step (a) are homologues to the segments of the sequence in step (b) as aligned in step (c), wherein identification of homology between the sequences of the first and third protein and the second and third protein identifies the first and second proteins as having a functional relationship.

Claim 37 (new): The method of claim 12 or claim 14, wherein the property of interest is a target for a drug.

Claim 38 (new): The method of claim 37, wherein the drug is an anti-microbial drug.

Claim 39 (new): The method of claim 12 or claim 14, wherein the property of interest is that of being essential for the growth or viability of an organism.

Claim 40 (new): The method of claim 12 or claim 14, wherein the first nucleic acid sequence or polypeptide sequence is derived from a pathogen.

Claim 41 (new): The method of claim 40, wherein the pathogen is a microorganism.

Claim 42 (new): The method of claim 41, wherein the microorganism is Mycobacterium tuberculosis (MTB).